Full wwPDB X-ray Structure Validation Report (i)

Apr 10, 2025 – 02:12 PM EDT

PDB ID	:	9CS1 / pdb_00009cs1
Title	:	E. coli BamA beta-barrel bound to darobactin and cyclic peptide CP2
Authors	:	Walker, M.E.; Gu, M.; Lu, J.; Klein, D.J.
Deposited on	:	2024-07-23
Resolution	:	2.15 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (i)) were used in the production of this report:

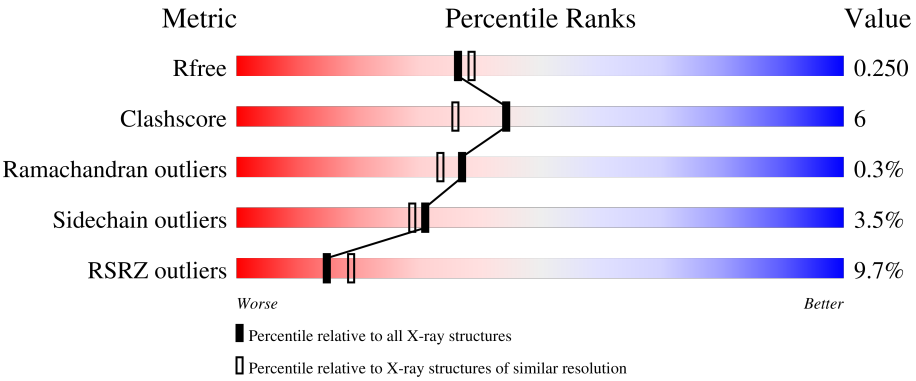
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	2.0rc1
EDS	:	3.0
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.006 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.42

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.15 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	164625	3336 (2.16-2.12)
Clashscore	180529	3585 (2.16-2.12)
Ramachandran outliers	177936	3554 (2.16-2.12)
Sidechain outliers	177891	3553 (2.16-2.12)
RSRZ outliers	164620	3337 (2.16-2.12)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	390	8% 81% 10% • 8%
2	C	17	12% 100%
3	D	7	29% 29% 14% 43% 14%

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 3145 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called Outer membrane protein assembly factor BamA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	360	Total	C	N	O	S	0	3	0
			2846	1816	457	564	9			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	421	MET	-	initiating methionine	UNP A7ZHR7

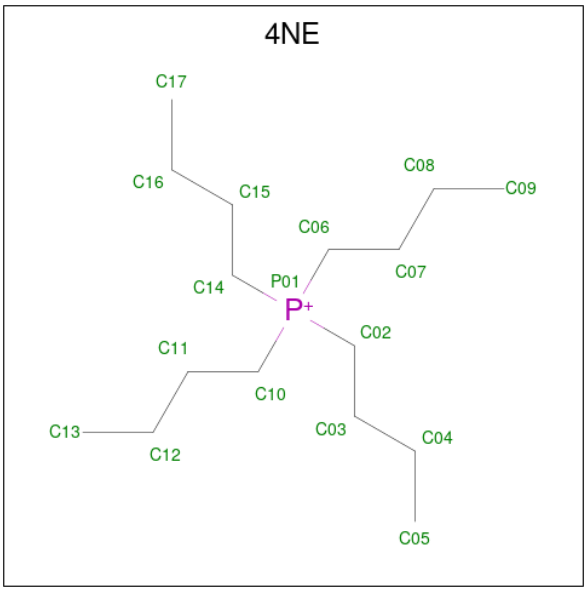
- Molecule 2 is a protein called Cyclic peptide CP2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	17	Total	C	N	O	S	0	0	1
			127	83	24	19	1			

- Molecule 3 is a protein called Darobactin A.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	D	7	Total	C	N	O	0	0	0
			70	47	11	12			

- Molecule 4 is tetrabutylphosphonium (CCD ID: 4NE) (formula: $C_{16}H_{36}P$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C P 17 16 1	0	0
4	A	1	Total C P 17 16 1	0	0

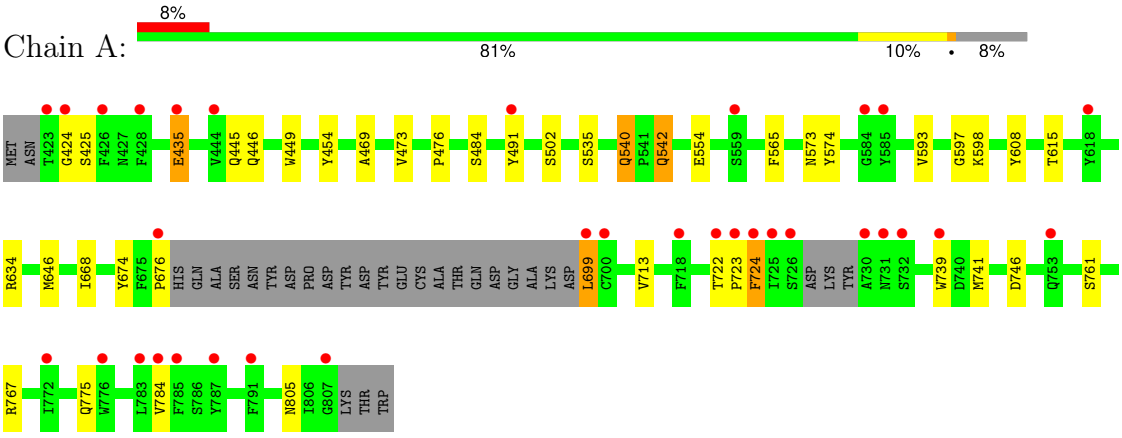
- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	59	Total O 59 59	0	0
5	C	9	Total O 9 9	0	0

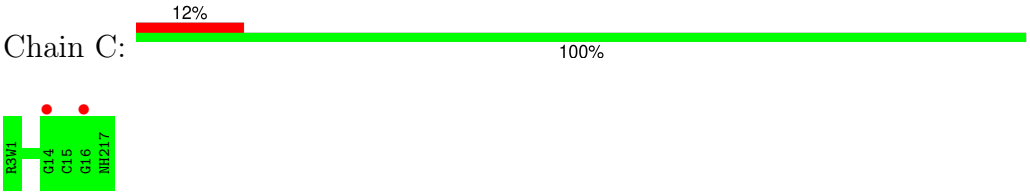
3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

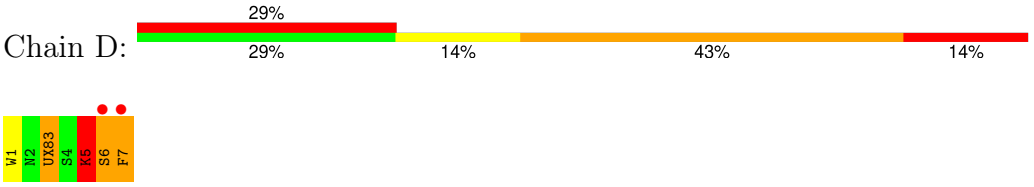
• Molecule 1: Outer membrane protein assembly factor BamA



• Molecule 2: Cyclic peptide CP2



• Molecule 3: Darobactin A



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 2 2 21	Depositor
Cell constants a, b, c, α, β, γ	58.60Å 82.62Å 127.89Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	26.64 – 2.15 26.64 – 2.15	Depositor EDS
% Data completeness (in resolution range)	83.0 (26.64-2.15) 82.9 (26.64-2.15)	Depositor EDS
R_{merge}	0.04	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.81 (at 2.15Å)	Xtriage
Refinement program	BUSTER 2.11.8 (22-FEB-2023)	Depositor
R, R_{free}	0.246 , 0.264 0.234 , 0.250	Depositor DCC
R_{free} test set	7381 reflections (5.11%)	wwPDB-VP
Wilson B-factor (Å^2)	52.8	Xtriage
Anisotropy	0.025	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 38.5	EDS
L-test for twinning[2]	$< \|L\| > = 0.51$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	3145	wwPDB-VP
Average B, all atoms (Å^2)	58.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.87% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: 4NE, NH2, UX8, DI8, R3W

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#$\|Z\|>5$	RMSZ	#$\|Z\|>5$
1	A	0.38	0/2937	0.67	0/3998
2	C	0.43	0/115	0.55	0/153
3	D	1.05	1/56 (1.8%)	0.90	0/71
All	All	0.40	1/3108 (0.0%)	0.67	0/4222

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	D	5	LYS	CA-CB	6.73	1.68	1.53

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2846	0	2615	32	0
2	C	127	0	107	0	0
3	D	70	0	50	13	0
4	A	34	0	72	0	0
5	A	59	0	0	1	0
5	C	9	0	0	0	0
All	All	3145	0	2844	35	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 6.

All (35) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:424:GLY:HA3	3:D:7:PHE:CE1	1.59	1.37
1:A:424:GLY:CA	3:D:7:PHE:CE1	2.27	1.18
1:A:424:GLY:HA3	3:D:7:PHE:HE1	0.94	0.99
1:A:424:GLY:CA	3:D:7:PHE:HE1	1.74	0.81
1:A:573[A]:ASN:ND2	5:A:1001:HOH:O	2.16	0.76
1:A:424:GLY:C	3:D:7:PHE:CE1	2.62	0.72
1:A:634:ARG:HB3	1:A:713:VAL:HB	1.81	0.61
1:A:775:GLN:HG2	1:A:784:VAL:HG22	1.85	0.56
3:D:3:UX8:CZ3	3:D:5:LYS:N	2.66	0.56
1:A:676:PRO:CG	1:A:699:LEU:HG	2.35	0.56
1:A:676:PRO:HG3	1:A:699:LEU:HG	1.88	0.54
1:A:540:GLN:HG2	1:A:674:TYR:CE1	2.47	0.50
1:A:574:TYR:CZ	1:A:597:GLY:HA3	2.48	0.48
1:A:723:PRO:O	1:A:724:PHE:HB2	2.14	0.48
1:A:554:GLU:HG3	1:A:565:PHE:CE1	2.50	0.46
1:A:668:ILE:HG23	1:A:767:ARG:HD3	1.98	0.46
1:A:739:TRP:CD1	1:A:741:MET:CE	2.99	0.46
1:A:435:GLU:H	1:A:435:GLU:HG2	1.60	0.45
1:A:469:ALA:HB3	1:A:491:TYR:CE2	2.52	0.45
1:A:425:SER:HA	3:D:6:SER:HA	1.99	0.45
1:A:746:ASP:O	1:A:761:SER:HA	2.17	0.44
1:A:454:TYR:CD2	1:A:473:VAL:HG12	2.53	0.44
1:A:424:GLY:HA3	3:D:7:PHE:CZ	2.36	0.43
1:A:424:GLY:N	3:D:7:PHE:CE1	2.85	0.43
3:D:3:UX8:CD2	3:D:3:UX8:C	2.96	0.43
1:A:739:TRP:CD1	1:A:741:MET:HE1	2.54	0.43
1:A:608:TYR:CD2	1:A:646:MET:HG3	2.54	0.43
1:A:476:PRO:HA	1:A:484:SER:HB3	2.00	0.43
3:D:1:TRP:CZ2	3:D:3:UX8:N	2.88	0.42
1:A:424:GLY:CA	3:D:7:PHE:CD1	2.96	0.41
1:A:446:GLN:HG2	1:A:449:TRP:HD1	1.85	0.41
1:A:425:SER:N	3:D:7:PHE:CZ	2.89	0.41
1:A:502:SER:OG	1:A:542:GLN:NE2	2.50	0.41
1:A:593:VAL:HG13	1:A:615:THR:HG22	2.02	0.41
1:A:722:THR:HA	1:A:723:PRO:HD3	1.92	0.40

There are no symmetry-related clashes.

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	A	357/390 (92%)	347 (97%)	9 (2%)	1 (0%)	37 33
2	C	13/17 (76%)	13 (100%)	0	0	100 100
3	D	4/7 (57%)	4 (100%)	0	0	100 100
All	All	374/414 (90%)	364 (97%)	9 (2%)	1 (0%)	37 33

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	724	PHE

5.3.2 Protein sidechains (i)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	A	303/326 (93%)	295 (97%)	8 (3%)	41 41
2	C	11/11 (100%)	11 (100%)	0	100 100
3	D	6/6 (100%)	3 (50%)	3 (50%)	0 0
All	All	320/343 (93%)	309 (97%)	11 (3%)	31 30

All (11) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	435	GLU
1	A	445	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	535	SER
1	A	540	GLN
1	A	542	GLN
1	A	598	LYS
1	A	699	LEU
1	A	805	ASN
3	D	5	LYS
3	D	6	SER
3	D	7	PHE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (6) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	427	ASN
1	A	445	GLN
1	A	459	ASN
1	A	542	GLN
1	A	555	HIS
3	D	2	ASN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains (i)

2 non-standard protein/DNA/RNA residues are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
3	UX8	D	3	3	14,16,17	1.43	2 (14%)	16,22,24	1.89	4 (25%)
2	DI8	C	7	2	11,13,14	0.45	0	13,17,19	0.89	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	UX8	D	3	3	-	0/5/10/12	0/2/2/2
2	DI8	C	7	2	-	0/1/11/13	0/2/2/2

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	D	3	UX8	CG-CD2	4.06	1.45	1.40
3	D	3	UX8	CG-CB	-2.80	1.48	1.52

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	D	3	UX8	CG-CB-CA	-5.97	103.74	111.58
3	D	3	UX8	CH2-CZ2-CE2	-2.39	116.82	120.09
3	D	3	UX8	O9-CB-CG	-2.29	107.41	111.08
3	D	3	UX8	O-C-CA	-2.01	119.59	124.77

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	D	3	UX8	3	0

5.5 Carbohydrates (i)

There are no oligosaccharides in this entry.

5.6 Ligand geometry (i)

2 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The

Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
4	4NE	A	901	-	16,16,16	0.22	0	18,18,18	0.49	0
4	4NE	A	902	-	16,16,16	0.38	0	18,18,18	0.38	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	4NE	A	901	-	-	4/20/20/20	-
4	4NE	A	902	-	-	7/20/20/20	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (11) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	902	4NE	C07-C06-P01-C10
4	A	902	4NE	C07-C06-P01-C02
4	A	902	4NE	C07-C06-P01-C14
4	A	902	4NE	C03-C02-P01-C10
4	A	902	4NE	C03-C02-P01-C06
4	A	902	4NE	C03-C02-P01-C14
4	A	901	4NE	C03-C02-P01-C10
4	A	901	4NE	C03-C02-P01-C06
4	A	901	4NE	C03-C02-P01-C14
4	A	901	4NE	C10-C11-C12-C13
4	A	902	4NE	C02-C03-C04-C05

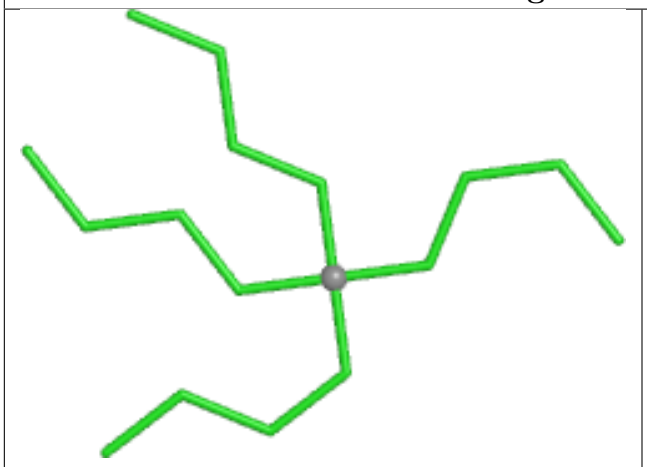
There are no ring outliers.

No monomer is involved in short contacts.

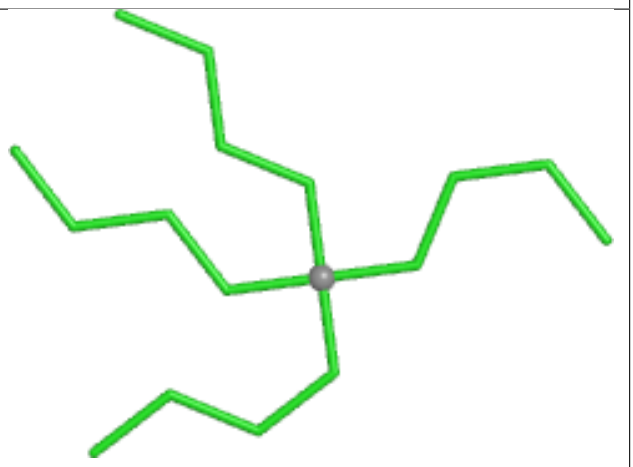
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will

also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

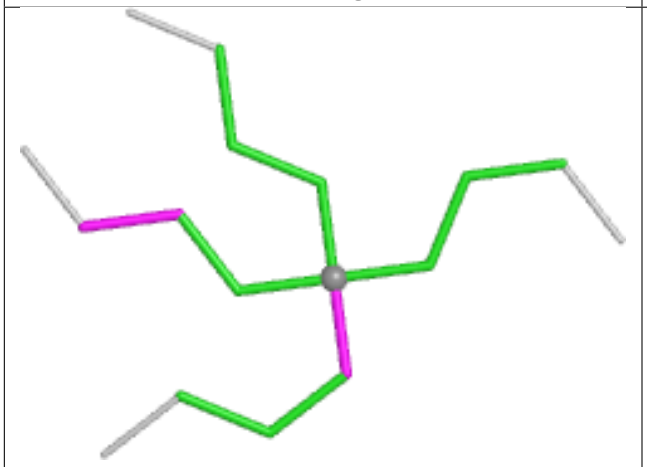
Ligand 4NE A 901



Bond lengths



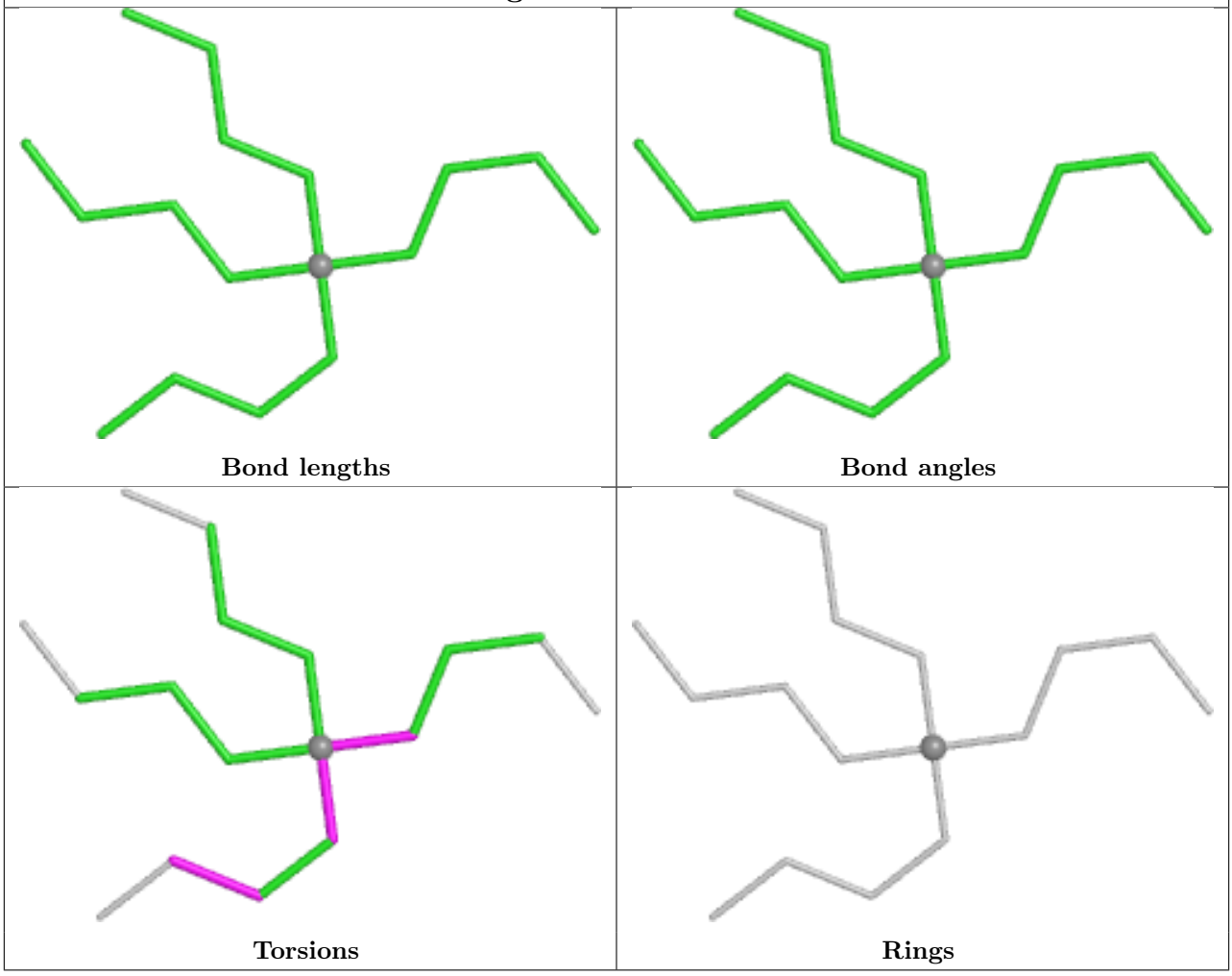
Bond angles



Torsions

Rings

Ligand 4NE A 902



5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95^{th} percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å^2)	Q<0.9
1	A	360/390 (92%)	0.72	33 (9%)	16	20	22, 56, 81, 116	3 (0%)
2	C	14/17 (82%)	0.16	2 (14%)	7	9	42, 44, 53, 59	0
3	D	6/7 (85%)	1.82	2 (33%)	1	2	63, 75, 83, 87	0
All	All	380/414 (91%)	0.72	37 (9%)	15	19	22, 56, 82, 116	3 (0%)

All (37) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	725	ILE	8.9
1	A	699	LEU	8.6
1	A	739	TRP	6.9
1	A	726	SER	5.9
1	A	730	ALA	5.1
1	A	724	PHE	4.9
1	A	718	PHE	4.6
1	A	491	TYR	3.9
1	A	732	SER	3.8
1	A	428	PHE	3.8
1	A	700	CYS	3.8
1	A	785	PHE	3.8
1	A	585	TYR	3.6
1	A	423	THR	3.6
3	D	7	PHE	3.6
1	A	776	TRP	3.1
1	A	807	GLY	3.0
2	C	16	GLY	2.9
1	A	791	PHE	2.7
1	A	676	PRO	2.6
1	A	783	LEU	2.6
1	A	424	GLY	2.5
2	C	14	GLY	2.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	753	GLN	2.5
1	A	435	GLU	2.3
1	A	618	TYR	2.3
1	A	444	VAL	2.3
1	A	731	ASN	2.3
1	A	722	THR	2.3
1	A	784	VAL	2.3
1	A	772[A]	ILE	2.3
1	A	723	PRO	2.2
3	D	6	SER	2.2
1	A	559	SER	2.1
1	A	584	GLY	2.1
1	A	426	PHE	2.1
1	A	787	TYR	2.0

6.2 Non-standard residues in protein, DNA, RNA chains (i)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
3	UX8	D	3	15/16	0.84	0.12	67,71,74,75	0
2	DI8	C	7	12/13	0.98	0.05	42,43,43,43	0

6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

6.4 Ligands (i)

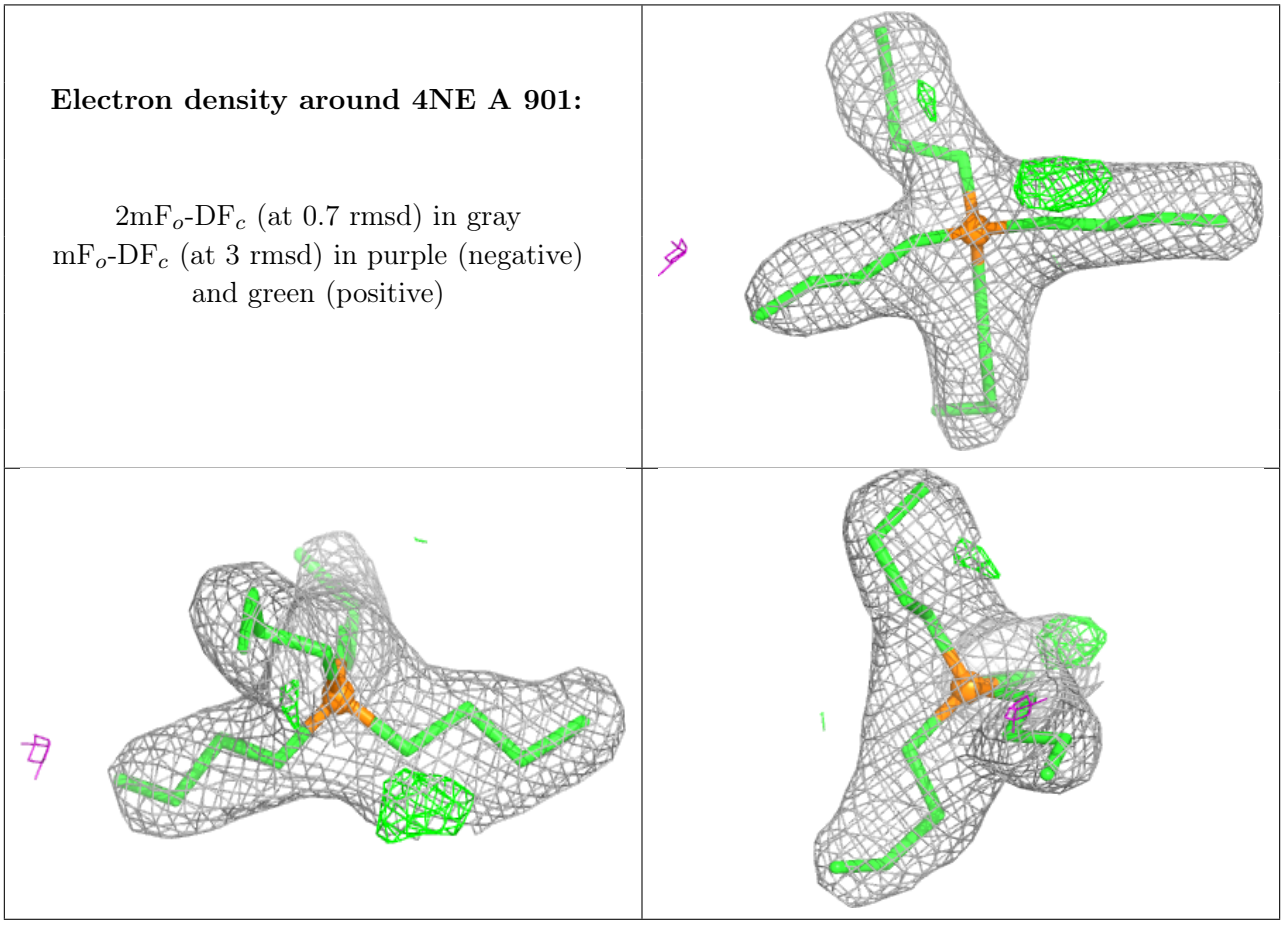
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	4NE	A	901	17/17	0.92	0.16	56,57,58,58	0
4	4NE	A	902	17/17	0.96	0.11	51,51,52,52	17

The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

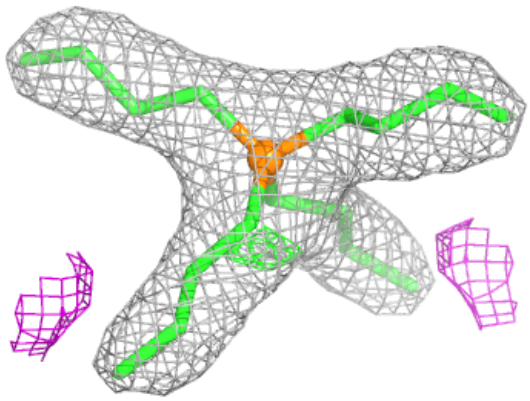
Electron density around 4NE A 901:

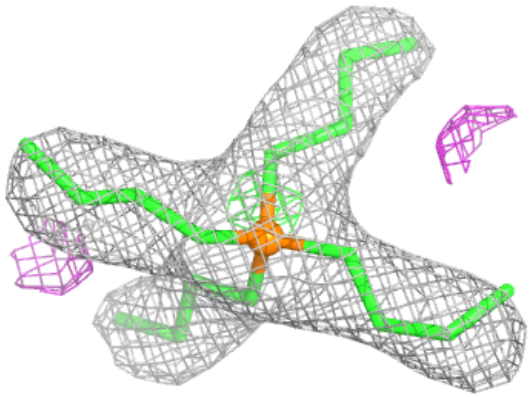
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

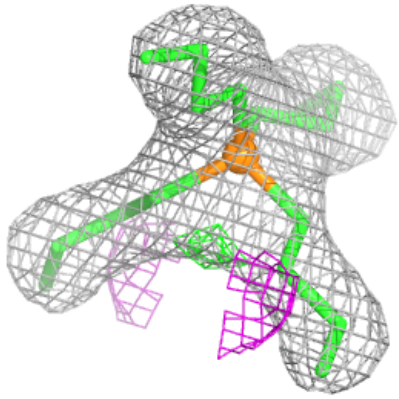


Electron density around 4NE A 902:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)







6.5 Other polymers

There are no such residues in this entry.